

School of Mathematical Sciences, Xiamen University

Numerical Solutions of Differential Equations

Chuanju XU

School of Mathematical Sciences
Xiamen University
361005 Xiamen
Fujian, China
email: cjxu@xmu.edu.cn

A course for 3rd year undergraduate students. The course covers 60 hours, and introduces some of the key methods used in the numerical solutions of various partial differential equations.



Contents

- Finite Difference methods
 - Ordinary differential equations
 - Elliptic equations
 - Parabolic equations
- Finite Element Methods for Elliptic Equations
 - Some functional spaces
 - Weak formulation
 - Galerkin methods
 - Finite Element Methods
- Finite Difference/Finite Element Methods for Parabolic Equations

Score composition: 50% continuous evaluation (homework, 2 computation practices) + 50% final exam



Goal: design and analysis of numerical methods for

$$u_t - \Delta u + g(u) = f.$$

Discretization:

Reducing the continuous problem to one with a finite number of unknowns

Basic alternatives:

- Replace derivatives with difference quotients (Finite Difference methods)
- Seek approximations in finite dimensional function spaces (Finite Element methods)

Numerical analysis:

- Understanding
- Error estimates
- Stability

Programming and Computing:

speed, use of memory, computational complexity



Partial differential operators

Second order linear partial differential operator:

$$\begin{aligned} Lu &= \sum_{i,j=1}^d D_i(a_{ij}D_j u) + \sum_{i=1}^d (D_i(b_i u) + c_i D_i u) + d_0 u \\ &= \nabla \cdot A \nabla u + \nabla \cdot (\mathbf{b} u) + \mathbf{c} \cdot \nabla u + d_0 u, \end{aligned}$$

where

- $u = u(\mathbf{x}) = u(x_1, x_2, \dots, x_d)$
- $D_i = \frac{\partial}{\partial x_i}, i = 1, \dots, d$
- $\nabla = (D_1, D_2, \dots, D_d)$
- The leading term $\nabla \cdot A \nabla u$ determines the type of the equation.



Classification of Equations

Let λ_i be the eigenvalues of A at point x .

- $\lambda_i \lambda_j > 0, \forall i, j$

\Rightarrow the equation is elliptic at x . Example:

$$u_{xx} + u_{yy} = f.$$

- $\lambda_i \neq 0$ and all but one λ_i have the same sign

\Rightarrow the equation is hyperbolic at x . Example:

$$u_{tt} - u_{xx} = f.$$

- There is at least one $\lambda_i = 0$

\Rightarrow the equation is parabolic at x . Example:

$$u_t - u_{xx} = f.$$



Finite Difference methods

$$u'(t) = \lim_{h \rightarrow 0} \frac{u(t+h) - u(t)}{h}.$$

Ordinary differential equations

Consider the initial value problem

$$\begin{aligned} u'(t) &= f(t, u(t)), \quad t \in (0, T] \\ u(0) &= u_0. \end{aligned}$$

For example,

$$\begin{cases} u'(t) = u(t) \tan(t) \\ u(0) = 1 \end{cases}$$



admits a solution $u(t) = \sec(t)$.

$$\begin{cases} u'(t) = \pi \cos(\pi t) \\ u(0) = 0 \end{cases}$$

admits a solution $u(t) = \sin(\pi t)$.

However, usually, no all IVPs allows a solution, or no analytical solutions are available.

Existence

For example

$$\begin{cases} u'(t) = 1 + u^2(t) \\ u(0) = 0 \end{cases}$$

admits a solution $u(t) = \tan t$ (local existence).



Theorem 2.1 If $f \in C^0(R)$ with $R = \{(t, u) : |t - t_0| \leq \alpha, |u - u_0| \leq \beta\}$. Then IVP has a solution $u(t)$ for $|t - t_0| \leq \min\{\alpha, \beta/M\}$, where $M = \max_{t \in R} |f(t, u(t))|$.

Uniqueness

Even if $f \in C^0$, no uniqueness is guaranteed!

Example 2.1

$$\begin{cases} u'(t) = u(t)^{2/3} \\ u(0) = 0 \end{cases}$$

have two solutions 0 and $\frac{1}{27}t^3$.

Theorem 2.2 If f and $\frac{\partial f}{\partial u} \in C^0(R)$. Then IVP has a unique solution in the interval $|t - t_0| \leq \min\{\alpha, \beta/M\}$.

Theorem 2.3 If f is continuous in the strip $a \leq t \leq b, -\infty < u < \infty$, and

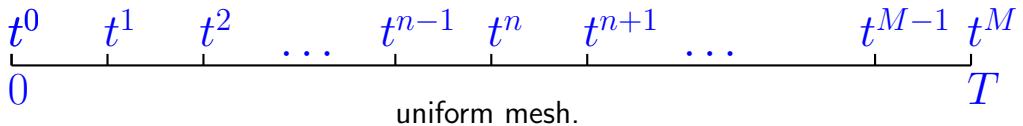
$$|f(t, u_1) - f(t, u_2)| \leq L|u_1 - u_2|$$



(Lipschitz condition). Then IVP has a unique solution in the interval $[a, b]$.

Finite difference methods

- Meshing: define a grid $t^n = nh, n = 0, 1, \dots, M; h = T/M$ on an interval $[0, T]$



- Question: $u(t^0) = u_0$ is known, how to compute $u(t^1), u(t^2), \dots, u(t^M)$
- Approximate derivatives with a difference quotient

$$u'(t^n) \simeq \frac{u(t^{n+1}) - u(t^n)}{h} \quad (\text{Forward difference})$$

$$u'(t^n) \simeq \frac{u(t^n) - u(t^{n-1})}{h} \quad (\text{Backward difference})$$



$$u'(t^{n+1/2}) \simeq \frac{u(t^{n+1}) - u(t^n)}{h} \quad (\text{Centered difference})$$

By using

$$u'(t^n) = f(t^n, u(t^n)), \quad \forall n = 0, 1, \dots, M$$

the above approximations lead to the following schemes

$$\frac{u^{n+1} - u^n}{h} = f(t^n, u^n), \quad n = 0, 1, \dots, M-1 \quad (\text{Forward Euler})$$

$$\frac{u^n - u^{n-1}}{h} = f(t^n, u^n), \quad n = 1, 2, \dots, M \quad (\text{Backward Euler})$$

$$\frac{u^{n+1} - u^n}{h} = \frac{f(t^{n+1}, u^{n+1}) + f(t^n, u^n)}{2}, \quad n = 0, 1, \dots, M-1 \quad (\text{Crank-Nicolson})$$

where u^n is an approximation of $u(t^n)$.



Remark 2.1 An alternative to Crank-Nicolson (trapezoidal) schema

$$\frac{u^{n+1} - u^{n-1}}{2h} = f(t^n, u^n), \quad n = 1, 2, \dots, M-1 \quad (\text{Leapfrog or Midpoint method})$$

which is a multistep method.

Another well-known multistep method is backward differentiation of second order:

$$\frac{3u^{n+1} - 4u^n + u^{n-1}}{2h} = f(t^{n+1}, u^{n+1}), \quad n = 1, 2, \dots, M-1 \quad (\text{BD2})$$

Explicit Adams-Basforth methods

Order-1: $u^{n+1} = u^n + hf^n$ (Forward Euler)



$$\text{Order-2: } u^{n+1} = u^n + h(3f^n - f^{n-1})/2$$

$$\text{Order-3: } u^{n+1} = u^n + h(23f^n - 16f^{n-1} + 5f^{n-2})/12$$

$$\text{Order-4: } u^{n+1} = u^n + h(55f^n - 59f^{n-1} + 37f^{n-2} - 9f^{n-3})/24$$

Implicit Adams-Moulton methods

$$\text{Order-2: } u^{n+1} = u^n + h(f^{n+1} + f^n)/2 \text{ (Crank-Nicolson schema)}$$

$$\text{Order-3: } u^{n+1} = u^n + h(5f^{n+1} + 8f^n - f^{n-1})/12$$

$$\text{Order-4: } u^{n+1} = u^n + h(9f^{n+1} + 19f^n - 5f^{n-1} + f^{n-2})/24$$

$$\text{Order-5: } u^{n+1} = u^n + h(251f^{n+1} + 646f^n - 264f^{n-1} + 106f^{n-2} - 19f^{n-3})/720$$



Analysis: accuracy and stability

1. Accuracy (error estimate)

- Truncation error: (Forward Euler)

$$u'(t^n) = \frac{u(t^{n+1}) - u(t^n)}{h} - R_f^n,$$

or

$$u(t^{n+1}) = u(t^n) + h u'(t^n) + r_f^n, \quad r_f^n = h R_f^n,$$

where r_f^n is the so-called *local truncation error*, which is the residual arising at the point t^{n+1} when we pretend that the exact solution “satisfies” the numerical schema. Estimation by using Taylor development:

$$R_f^n = \frac{h}{2} u''(\xi^n), \quad r_f^n = \frac{h^2}{2} u''(\xi^n), \quad \xi^n \in [t^n, t^{n+1}].$$



Similarly, the local truncation error is

$$r_b^n = \frac{h^2}{2} u''(\xi^n), \quad \xi^n \in [t^{n-1}, t^n]$$

for Backward Euler, and

$$r_c^n = -\frac{h^3}{48} (u'''(\xi^n) + u'''(\tilde{\xi}^n)), \quad \xi^n \in [t^{n+1/2}, t^{n+1}], \quad \tilde{\xi}^n \in [t^n, t^{n+1/2}]$$

for Centered Euler.

r^n = local error $u(t^{n+1}) - u^{n+1}$, if the previous solutions u^n, u^{n-1}, \dots are exact.

- Error equations

Let $e^n = u(t^n) - u^n, n = 0, 1, \dots, M$, then it holds

$$\frac{e^{n+1} - e^n}{h} + R_f^n = f(t^n, u(t^n)) - f(t^n, u^n) \quad (\text{Forward Euler})$$



$$\frac{e^n - e^{n-1}}{h} + R_b^n = f(t^n, u(t^n)) - f(t^n, u^n) \quad (\text{Backward Euler})$$

$$\frac{e^{n+1} - e^n}{h} + R_c^n = \frac{f(t^{n+1}, u(t^{n+1})) + f(t^n, u(t^n))}{2} - \frac{f(t^{n+1}, u^{n+1}) + f(t^n, u^n)}{2}$$

(Crank-Nicolson)

Hypothesis: f is Lipschitz continuous w.r.t. the second variable, i.e.,

$$|f(t, u) - f(t, v)| \leq L|u - v|, \quad \forall t \in [0, T], \forall u, v \in R.$$

- Forward Euler



Let $R = \max_n |R_f^n|$. Then

$$\begin{aligned}|e^{n+1}| &\leq |e^n| + hL|e^n| + hR \\&\leq (1 + hL)^{n+1}|e^0| + \frac{hR}{hL}[(1 + hL)^{n+1} - 1] \\&\leq e^{LT}|e^0| + \frac{R}{L}(e^{LT} - 1) \\&\leq O(h), \quad n = 0, 1, \dots, M-1.\end{aligned}$$

Exercise 2.1 Carry out an error analysis for Backward Euler schema.

Exercise 2.2 Carry out an error analysis for Crank-Nicolson (Trapezoidal) schema.

2. Stability

Convergence guaranteed as the grid is refined ($h \rightarrow 0$), but in practice a calculation has to be performed with some nonzero time step.



Question: can h be chosen only with accuracy consideration?

Examples of unstable computations

Example 2.2 Consider the IVP:

$$\frac{du(t)}{dt} = \cos t, \quad u(0) = 0$$

with solution $u(t) = \sin t$.

- Forward Euler is used to solve this problem up to $T = 1$.
- The error R is

$$R(t) = \frac{1}{2}hu''(t) + O(h^2) = -\frac{1}{2}h \sin t + O(h^2).$$



- Since $f(t) = \sin t$ is independent of u , it is Lipschitz continuous w.r.t. u with Lipschitz constant $L = 0$
- the error estimate

$$|e^{n+1}| \leq |e^n| + hR \leq |e^0| + nhR \leq T \frac{1}{2}h \max_t |\sin t| + O(h^2) \leq \frac{h}{2} + O(h^2).$$

If we want to compute a solution with an error $\leq 10^{-2}$. Then we should take $h \leq 2 \times 10^{-2}$ and time steps $T/h = 50$. Indeed, calculating using $h = 2 \times 10^{-2}$ gives a numerical solution $u^{50} = 0.84603991$ with an error $e^{50} = \sin 1 - u^{50} = -0.45689 \times 10^{-2}$.

Example 2.3 Now suppose we modify the above equation to

$$\frac{du(t)}{dt} = \lambda(u(t) - \sin t) + \cos t, \quad u(0) = 0, \quad (1)$$

where λ is a constant, say $\lambda = -10$. The solution is the same as before, $u(t) = \sin t$.



- $h = ?$ to get an error $\leq 10^{-2}$.
- Since the local truncation error depends only on the true solution $u(t)$, which is unchanged from Example 2.2, we might hope that we could use the same h as in that example, $h = 2 \times 10^{-2}$.
- In fact $h = 2 \times 10^{-2}$ gives $u^{50} = 0.84225545$ with an error $e^{50} = -0.78446 \times 10^{-3}$.
- the error is even smaller than in Example 2.2.

Example 2.4 Now consider the problem (1) with $\lambda = -200$ and the same data as before.

- The solution is unchanged and so is the local truncation error.
- Computation with the same step size: $h = 2 \times 10^{-2} \Rightarrow u^{50} = -0.5983 \times 10^{17}$.
 \Rightarrow Computation is unstable, and the error grows exponentially in time.
- The method is convergent, and indeed with sufficiently small time steps we obtain very good results, as shown in Table 1.



- Something happens between the values $h = 0.0125$ and $h = 0.0100$.
- For smaller values of h we get very good results, whereas for larger values of h we lost accuracy.
- The global error satisfies

$$e^{n+1} = (1 + h\lambda)e^n + hR^n.$$

\Rightarrow source of the exponential growth in the error — in each time step the previous error is multiplied by a factor of $1 + h\lambda$.

- For the case $\lambda = -200$ and $h = 1.25 \times 10^{-2}$, we have $1 + h\lambda = -1.5$ and After 50 steps we expect the error introduced in the first step to have grown by a factor of roughly $(1.5)^{50} \simeq 10^7$.
- In Example 2.3 with $\lambda = -10$, $h = 0.02$, we have $1 + h\lambda = 0.8$, causing a decay in the effect of previous errors in each step.
- This explains why we got a better result in Example 2.3 than in Example 2.2 where $1 + h\lambda = 1$.



<i>h</i>	Errors
0.02000	0.59830516E+17
0.01250	-0.31839689E+07
0.01000	-0.21056104E-04
0.00100	-0.20973301E-05

Table 1: Errors as a function of *h*.

3. Absolute stability

Model problem

$$\frac{du}{dt} = \lambda u,$$

where λ is a constant.

- Forward Euler



$$u^{n+1} = (1 + h\lambda)u^n.$$

Taking into account the round-off errors, we will obtain $\{\bar{u}^n\}$ rather than $\{u^n\}$

$$\bar{u}^{n+1} = \bar{u}^n + h\lambda\bar{u}^n + \varepsilon^n.$$

Error $\delta^n := \bar{u}^n - u^n$ satisfies

$$\delta^{n+1} = \delta^n + h\lambda\delta^n + \varepsilon^n.$$

\Rightarrow

$$\delta^{n+1} = (1 + h\lambda)^{n+1}\delta^0 + (1 + h\lambda)^n\varepsilon^0 + \cdots + \varepsilon^n.$$

Suppose $\varepsilon^n \leq \varepsilon$, then

$$|\delta^{n+1}| \leq |1 + h\lambda|^{n+1}|\delta^0| + \varepsilon \frac{|1 + h\lambda|^{n+1} + 1}{-h\lambda} \rightarrow c\varepsilon \quad \text{if } |1 + h\lambda| < 1.$$



Notion This schema is absolutely stable when $|1 + h\lambda| < 1$; otherwise it is unstable.

- ★ There are two parameters h and λ , but only their product $z = h\lambda$ matters.
- ★ The method is stable whenever $-2 \leq z \leq 0$, and we say that the **Interval of Absolute Stability** for Forward Euler method is $[-2, 0]$.
- ★ **Absolute Stability Region:** region in the complex plane, defined as the set of the complex number z such that the amplitude coefficient is smaller than one. That is, allowing complex λ , let $z = h\lambda$:

$$\text{ASR(F.E.)} = \{z : |1 + z| \leq 1\}.$$

$$\text{ASR(B.E.)} = \{z : \frac{1}{|1-z|} \leq 1\}.$$

...

For multi-step schemes, say Leapfrog,

$$\begin{pmatrix} u^{n+1} \\ u^n \end{pmatrix} = \begin{pmatrix} 2h\lambda & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u^n \\ u^{n-1} \end{pmatrix}.$$



$$\left\| \begin{pmatrix} u^{n+1} \\ u^n \end{pmatrix} \right\| \leq \left\| \begin{pmatrix} 2h\lambda & 1 \\ 1 & 0 \end{pmatrix} \right\| \left\| \begin{pmatrix} u^n \\ u^{n-1} \end{pmatrix} \right\|,$$

where $\|v^n\|$ is a vector norm, $\|M\|$ is the subordinate matrix norm (refer to vector and matrix norm.pdf).

$$\text{ASR}(Lf) \supseteq \{z : \left\| \begin{pmatrix} 2z & 1 \\ 1 & 0 \end{pmatrix} \right\| \leq 1\}.$$

Remark 2.2 More precise definition of ASR is the set of z such that the scheme produces bounded solutions; see, e.g., [Finite Difference Methods for Ordinary and Partial Differential Equations [LeVeque1955].pdf, chapter 7, p154].

Exercise 2.3 Determine the absolute stability region of the Crank-Nicolson schema and Leapfrog schema.

A one-step method for solving $u' = \lambda u$ (Dahlquist test equation, 1963) can be described as

$$u^{n+1} = g(z)u^n, n = 0, 1, \dots, M-1,$$



where $g(z) = P(z)/Q(z)$ is a rational function and $P(z)$ and $Q(z)$ are polynomials.

The set

$$S = \{z \in \mathbb{C} : |g(z)| \leq 1\}$$

is called the absolute stability domain of the method.

* A method whose absolute stability domain satisfies

$$S \supset \mathbb{C}^- := \{z : \operatorname{Re} z < 0\}$$

is called A-stable, where \mathbb{C}^- denotes the entire left half-plane.

* Complex λ comes from solving a system of ODEs:

$$\frac{d\mathbf{u}}{dt} = A\mathbf{u},$$

where λ is an eigenvalue of A .



★ Suppose A is diagonalizable: $A = T^{-1}\Lambda T$, with $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_N\}$, then

$$\frac{d\mathbf{v}}{dt} = \Lambda\mathbf{v},$$

where $\mathbf{v} = T\mathbf{u}$.

Exercise 2.4

$$\begin{cases} \frac{d\mathbf{u}}{dt} = -A\mathbf{u}, & t > 0, \\ \mathbf{u}(0) = (1, 1)^T, \end{cases}$$

where

$$A = \begin{pmatrix} 99 & 7\sqrt{2} \\ 7\sqrt{2} & 2 \end{pmatrix}.$$

Solve numerically the problem and investigate the stability and convergence.



Exercise 2.5 Solve numerically the problem:

$$\begin{cases} \frac{d\mathbf{u}}{dt} = A\mathbf{u}, & 0 < t \leq 1, \\ \mathbf{u}(0) = (0, 2)^T, \end{cases}$$

where

$$A = \begin{pmatrix} -50 & 49 \\ 49 & -50 \end{pmatrix}.$$

Compare the numerical solution with the exact solution

$$\mathbf{u} = (\exp(-t) - \exp(-99t), \exp(-t) + \exp(-99t))^T.$$

Hint: Solving the characteristic equation $|A - \lambda I| = 0$ gives two eigenvalues $\lambda_1 = -1, \lambda_2 = -99$. Furthermore, it can be checked that

$$A = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & -99 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$



Let

$$C = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad \Lambda = \begin{pmatrix} -1 & 0 \\ 0 & -99 \end{pmatrix},$$

then

$$\frac{d\mathbf{u}}{dt} = \frac{1}{2} C^T \Lambda C \mathbf{u}.$$

Let $\mathbf{v} = C\mathbf{u}$, then

$$\frac{d\mathbf{v}}{dt} = \Lambda \mathbf{v},$$

since $(\frac{1}{2}C^T)^{-1} = C$.

Obviously, $\mathbf{v} = (2 \exp(-t), 2 \exp(-99t))^T$ is the solution of

$$\begin{cases} \frac{d\mathbf{v}}{dt} = \Lambda \mathbf{v}, & t > 0, \\ \mathbf{v}(0) = (2, 2)^T = C(0, 2)^T. \end{cases}$$



Thus,

$$\mathbf{u} = C^{-1}\mathbf{v} = \frac{1}{2}C^T\mathbf{v} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 \exp(-t) \\ 2 \exp(-99t) \end{pmatrix}$$

is the solution of the original problem.

Exercise 2.6 Solve numerically by several schemes the following problem:

$$\begin{cases} \frac{d\mathbf{u}}{dt} = A\mathbf{u}, & 0 < t \leq 1, \\ \mathbf{u}(0) = \mathbf{u}_0, \end{cases}$$



where

$$A = -(N+1)^2 \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 2 & -1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & & & & \\ & & & \ddots & & & \\ 0 & \cdots & 0 & 0 & -1 & 2 & -1 \\ 0 & \cdots & 0 & 0 & 0 & -1 & 2 \end{pmatrix}_{N \times N}. \quad (2)$$

$$\mathbf{u}_0 = \begin{pmatrix} \frac{1}{N+1} \\ \frac{2}{N+1} \\ \vdots \\ \frac{N}{N+1} \end{pmatrix}.$$

Contents of the report:



Title: Numerical investigation of a number of schemes

Abstract (the goal: investigating the properties of different schemes: accuracy, stability, and computational complexity through numerical experiments)

1) Description of the schemes: forward Euler, backward Euler, Central, leapfrog, BD2, etc. to the ODE problem:

$$\begin{aligned} u' &= f(t, u), \quad t \in (0, T], \\ u(0) &= u_0. \end{aligned}$$

2) Analysis of these schemes (indicate the known results about the truncation errors and convergence order, stability, and computational complexity)

3) Numerical examples: computation configuration (equation, initial condition, domain, mesh), results and interpretation (via tables and figures)

4) Conclusion

5) Appendix: code



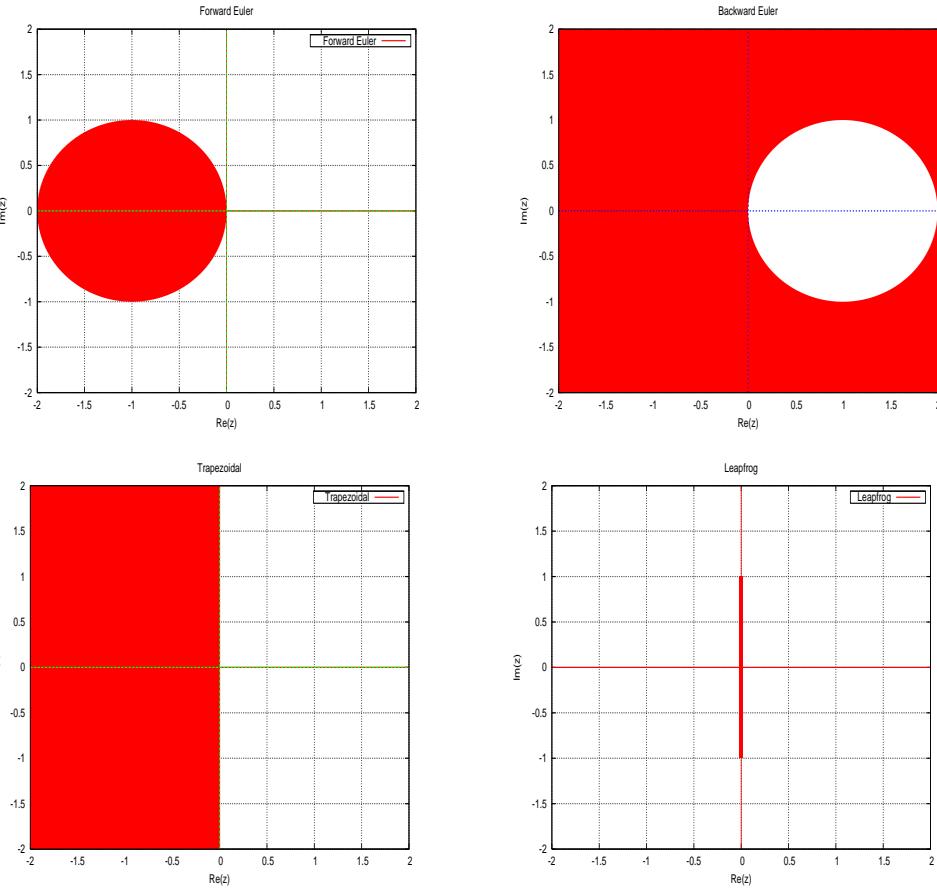


Figure 1: Stability regions.



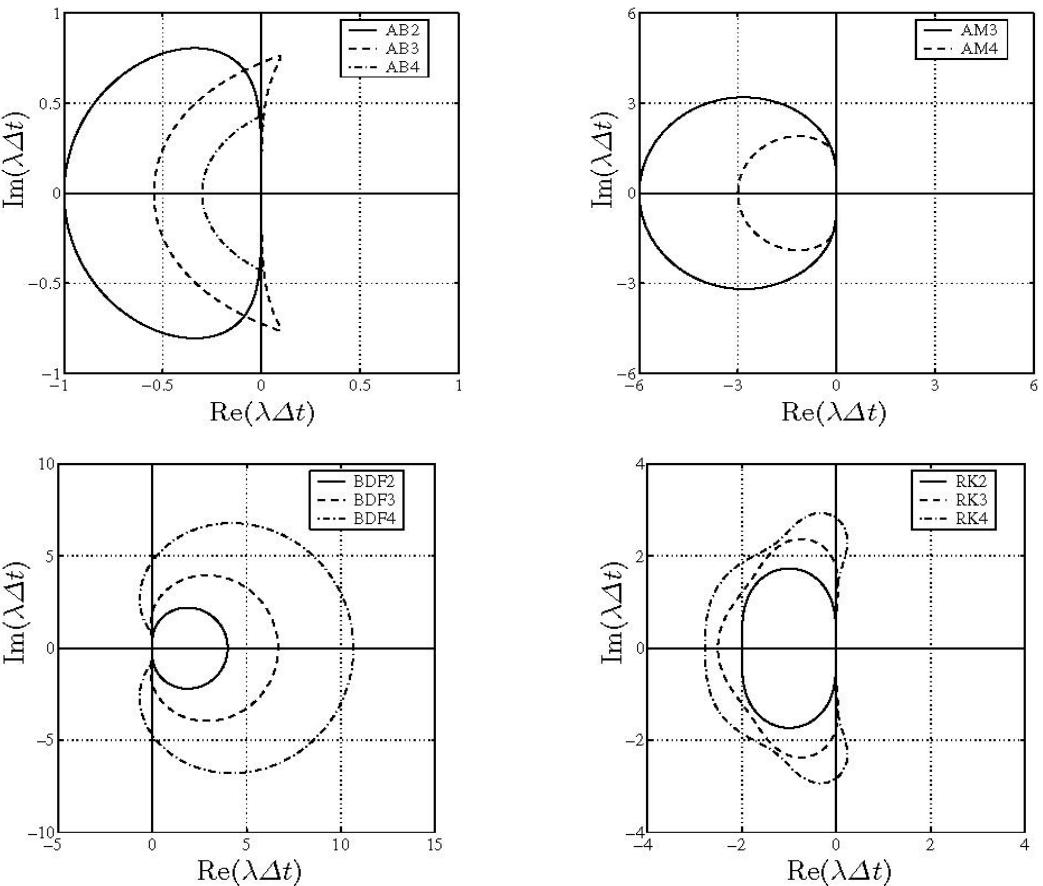


Figure 2: Stability regions of AB, AM, BDF and RK methods. Note AS regions of BDF are outside parts.

Elliptic equations

Let's consider the following problem

$$\begin{cases} Lu(x) = f(x), \quad \forall x \in (a, b) \\ u(a) = u(b) = 0, \end{cases}$$

where L is a linear elliptic operator.

Example:

$$Lu = -u''.$$

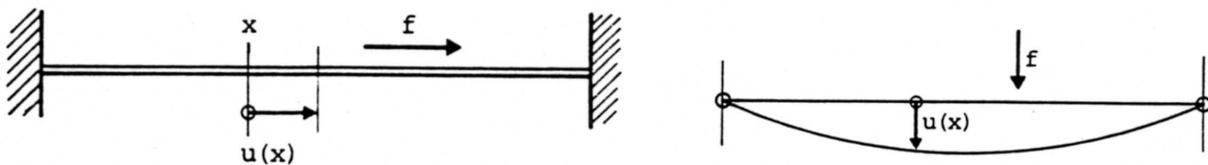
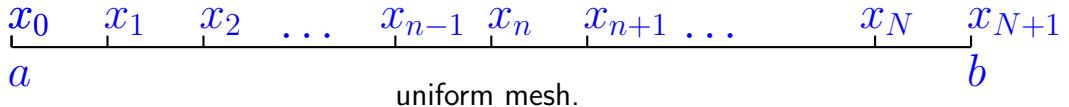


Figure 3: An elastic bar (left) and an elastic cord (right).

Let $\{x_i\}_{i=0}^{N+1}$, $x_i = a + ih$, $h = \frac{b-a}{N+1}$, be a grid in the interval $[a, b]$:



uniform mesh.

Let L_h be a discrete operator, approximating L .

An example:

$$Lu = -\frac{\partial^2 u}{\partial x^2}, \quad L_h u_i = -\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2},$$

then

$$L_h u(x_i) \simeq [Lu](x_i).$$

Construction of a schema

Note that

$$[Lu](x_i) = f(x_i), \quad i = 1, 2, \dots, N$$



can be approximated by

$$L_h u(x_i) \simeq f(x_i), \quad i = 1, 2, \dots, N.$$

This leads to the schema

$$L_h u_i = f(x_i), \quad i = 1, 2, \dots, N$$

subject to the boundary condition $u_0 = u_{N+1} = 0$.

Example:

$$-\frac{\partial^2 u}{\partial x^2}(x_i) = f(x_i), \quad i = 1, 2, \dots, N$$

is approximated by

$$-\frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1})}{h^2} \simeq f(x_i), \quad i = 1, 2, \dots, N.$$



This suggests

$$-\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} = f(x_i), \quad i = 1, 2, \dots, N,$$

together with $u_0 = u_{N+1} = 0$.

- * This schema is called central schema.
- * If we set $a = 0, b = 1$, then the coefficient matrix is exactly the same as the one defined in (2), i.e.,

$$A = (N+1)^2 \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & & & & \\ & \cdot & \cdot & \cdot & & & \\ 0 & \cdots & 0 & 0 & -1 & 2 & -1 \\ 0 & \cdots & 0 & 0 & 0 & -1 & 2 \end{pmatrix}_{N \times N}. \quad (3)$$



Error analysis

- Truncation error: $R_i = L_h u(x_i) - [Lu](x_i)$ or equivalently $R_i = L_h u(x_i) - f_i$.
- Consistency: the difference schema is consistent with the original problem if $R_i \rightarrow 0$ as $h \rightarrow 0$ for all $i = 1, 2, \dots, N$.
- Let U_1 and U_2 be two approximate solutions with two different RHS functions f_1 and f_2 . The difference schema is stable, if

$$\|U_1 - U_2\| \leq c \|f_1 - f_2\|,$$

or equivalently

$$\|U\| \leq c \|f\|$$

if L_h is linear.

- A difference schema is convergent, if for any

$$u_j \rightarrow u(x_j), \quad \forall j = 0, 1, \dots, N + 1$$



as the computational grid is refined, i.e. $h \rightarrow 0$.

Example:

$$Lu = -u'', \quad L_h u_i = -\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2}.$$

Then L_h is consistent, because

$$R_i = L_h u(x_i) - [Lu](x_i) = O(h^2) \rightarrow 0, \text{ as } h \rightarrow 0.$$

Stability, i.e., energy inequality in the continuous case: For $Lu = -u''$, there exists a constant c , such that

$$\|u\|_0 \leq c\|f\|_0, \quad \|u\|_1 \leq c\|f\|_0.$$

Exercise 2.7 Prove that the consistency and stability leads to the convergence.



Exercise 2.8 Consider the variable coefficient equation

$$Lu = f(x), \quad u(a) = u(b) = 0,$$

where $Lu = -(pu')'(x)$, $p_M \geq p(x) \geq p_0 > 0, \forall x \in [a, b]$.

1) Establish an energy inequality.

2) Set $p(x) \equiv 1$. Consider the central schema on the non-uniform mesh $\{x_i\}_{i=0}^{N+1}, h_i = x_i - x_{i-1}$:

$$\begin{cases} L_h u_i = f(x_i), & i = 1, 2, \dots, N, \\ u_0 = u_{N+1} = 0, \end{cases}$$

where

$$L_h u_i = -\frac{\frac{u_{i+1}-u_i}{h_{i+1}} - \frac{u_i-u_{i-1}}{h_i}}{\frac{h_i+h_{i+1}}{2}}, \quad i = 1, 2, \dots, N.$$

Analyze the truncation error $R_i = L_h[u(x_i)] - [Lu](x_i)$, $i = 1, 2, \dots, N$ in term of $h = \max_{1 \leq i \leq N+1} |h_i|$.



Exercise 2.9 Derive some energy inequalities for $Lu = -(pu')' + u'$, $u(a) = u(b) = 0$, where $p_M \geq p(x) \geq p_0 > 0$, $\forall x \in [a, b]$.

Stability of the discrete problem

Some notations. Define

$$I_h = \{x_1, \dots, x_N\}, \quad \bar{I}_h = \{x_0, x_1, \dots, x_{N+1}\}, \quad I_h^+ = \{x_1, \dots, x_{N+1}\}.$$

$$v_h = \{v_0, v_1, \dots, v_{N+1}\}$$

is a discrete function defined in \bar{I}_h .

Difference operators

$$(v_i)_{\bar{x}} = v_{i,\bar{x}} = \frac{v_i - v_{i-1}}{h_i}, \quad (v_i)_x = v_{i,x} = \frac{v_{i+1} - v_i}{h_{i+1}},$$



$$(v_i)_{\hat{x}} = v_{i,\hat{x}} = \frac{v_{i+1} - v_i}{\bar{h}_i}, \text{ with } \bar{h}_i = \frac{1}{2}(h_i + h_{i+1}), \bar{h}_0 = \frac{1}{2}h_1, \bar{h}_{N+1} = \frac{1}{2}h_{N+1}.$$

Inner product

$$(u_h, v_h)_{I_h} = \sum_{I_h} u_i v_i \bar{h}_i, \quad (u_h, v_h)_{\bar{I}_h} = \sum_{\bar{I}_h} u_i v_i \bar{h}_i, \quad (u_h, v_h)_{I_h^+} = \sum_{I_h^+} u_i v_i h_i.$$

Norms:

$$\|v_h\|_c = \max_{\bar{I}_h} |v_i|, \quad \|v_h\|_0 = (v_h, v_h)_{\bar{I}_h}^{1/2},$$

$$|v_h|_1 = ((v_h)_{\bar{x}}, (v_h)_{\bar{x}})_{I_h^+}^{1/2}, \quad \|v_h\|_1^2 = \|v_h\|_0^2 + |v_h|_1^2.$$

Relation

$$(v_i w_i)_{\bar{x}} = (v_i)_{\bar{x}} w_{i-1} + v_i (w_i)_{\bar{x}}. \quad (4)$$



Exercise 2.10 Prove (4).

Thus

$$\sum_{i=m+1}^n (v_i w_i)_{\bar{x}} h_i = \sum_{i=m+1}^n (v_i)_{\bar{x}} w_{i-1} h_i + \sum_{i=m+1}^n v_i (w_i)_{\bar{x}} h_i.$$

First term of RHS:

$$\sum_{i=m+1}^n (v_i)_{\bar{x}} w_{i-1} h_i = \sum_{i=m+1}^n (v_i - v_{i-1}) w_{i-1} = \sum_{i=m}^{n-1} (v_{i+1} - v_i) w_i = \sum_{i=m}^{n-1} (v_i)_x w_i h_{i+1}.$$

LHS:

$$\sum_{i=m+1}^n (v_i w_i)_{\bar{x}} h_i = \sum_{i=m+1}^n (v_i w_i - v_{i-1} w_{i-1}) = v_n w_n - v_m w_m.$$



This leads to Discrete Integral by Part:

$$\sum_{i=m+1}^n v_i(w_i)_{\bar{x}} h_i = - \sum_{i=m}^{n-1} (v_i)_x w_i h_{i+1} + v_n w_n - v_m w_m.$$

or in an alternative form (using $(v_m)_x w_m h_{m+1} = v_{m+1} w_m - v_m w_m$)

$$\sum_{i=m+1}^{n-1} (v_i)_x w_i h_{i+1} = - \sum_{i=m+1}^n v_i(w_i)_{\bar{x}} h_i + v_n w_n - v_{m+1} w_m. \quad (5)$$

Exercise 2.11 Prove (5).

Thanks to

$$(v_i)_x h_{i+1} = (v_i)_{\hat{x}} \bar{h}_i$$



(5) becomes

$$\sum_{i=m+1}^{n-1} (v_i)_{\hat{x}} w_i \bar{h}_i = - \sum_{i=m+1}^n v_i (w_i)_{\bar{x}} h_i + v_n w_n - v_{m+1} w_m.$$

Taking $v_i = (u_i)_{\bar{x}}$, $w_i = v_i$, then (Difference Green Formula)

$$\begin{aligned} \sum_{i=m+1}^{n-1} ((u_i)_{\bar{x}})_{\hat{x}} v_i \bar{h}_i &= - \sum_{i=m+1}^n (u_i)_{\bar{x}} (v_i)_{\bar{x}} h_i + (u_n)_{\bar{x}} v_n - (u_{m+1})_{\bar{x}} v_m \\ &= - \sum_{i=m+1}^n (u_i)_{\bar{x}} (v_i)_{\bar{x}} h_i + (u_n)_{\bar{x}} v_n - (u_m)_x v_m. \end{aligned}$$

Particular case: $m = 0, n = N + 1$

$$(((u_h)_{\bar{x}})_{\hat{x}}, v_h)_{I_h} = -((u_h)_{\bar{x}}, (v_h)_{\bar{x}})_{I_h^+} + (u_{N+1})_{\bar{x}} v_{N+1} - (u_0)_x v_0.$$

Remark 2.3 It holds a more general Green Formula.



Taking $v_i = p_{i-1/2}(u_i)_{\bar{x}}, w_i = v_i$, then

$$\begin{aligned} \sum_{i=m+1}^{n-1} (p_{i-1/2}(u_i)_{\bar{x}})_{\hat{x}} v_i \bar{h}_i &= - \sum_{i=m+1}^n p_{i-1/2}(u_i)_{\bar{x}} (v_i)_{\bar{x}} h_i + p_{n-1/2}(u_n)_{\bar{x}} v_n - p_{m+1/2}(u_{m+1})_{\bar{x}} v_m \\ &= - \sum_{i=m+1}^n p_{i-1/2}(u_i)_{\bar{x}} (v_i)_{\bar{x}} h_i + p_{n-1/2}(u_n)_{\bar{x}} v_n - p_{m+1/2}(u_m)_x v_m. \end{aligned}$$

If $m = 0, n = N + 1$

$$((p_h(u_h)_{\bar{x}})_{\hat{x}}, v_h)_{I_h} = - (p_h(u_h)_{\bar{x}}, (v_h)_{\bar{x}})_{I_h^+} + p_{N+1/2}(u_{N+1})_{\bar{x}} v_{N+1} - p_{1/2}(u_0)_x v_0.$$

Some inequalities



Cauchy inequality: if the matrix (α_{ij}) is symmetric positive,

$$\left| \sum_{i,j=0}^{N+1} \alpha_{ij} a_i b_j \right| \leq \left(\sum_{i,j=0}^{N+1} \alpha_{ij} a_i a_j \right)^{1/2} \left(\sum_{i,j=0}^{N+1} \alpha_{ij} b_i b_j \right)^{1/2}.$$

Particular case: if $\alpha_{ij} = \bar{h}_i \delta_{ij}$, then (Discrete Schwarz Inequality)

$$|(u_h, v_h)_{\bar{I}_h}| \leq (u_h, u_h)_{\bar{I}_h}^{1/2} (v_h, v_h)_{\bar{I}_h}^{1/2}.$$

Lemma 2.1 (Discrete Poincaré Inequality) v_h is a discrete function defined in \bar{I}_h , and $v_0 = v_{N+1} = 0$. Then

$$\|v_h\|_c^2 \leq \frac{b-a}{4} |v_h|_1^2.$$

PROOF. Note that

$$v_i = \sum_{j=1}^i v_{j,\bar{x}} h_j, \quad v_i = - \sum_{j=i+1}^{N+1} v_{j,\bar{x}} h_j.$$



By using Cauchy inequality to the above equalities, we have

$$v_i^2 \leq (x_i - a) \sum_{j=1}^i v_{j,\bar{x}}^2 h_j, \quad v_i^2 \leq (b - x_i) \sum_{j=i+1}^{N+1} v_{j,\bar{x}}^2 h_j.$$

Multiplying resp. $(b-x_i)$ and (x_i-a) to the above two inequalities, and summing the resulting inequalities give

$$v_i^2 \leq \frac{(x_i - a)(b - x_i)}{b - a} |v_h|_1^2 \leq \frac{b - a}{4} |v_h|_1^2, \quad \forall i \in \bar{I}_h.$$

Remark 2.4

- 1) *Discrete Poincaré Inequality, i.e. Lemma 2.1, still holds if only $v_0 = 0$ or $v_{N+1} = 0$.*
- 2) *Under same assumption as in Lemma 2.1, it holds*

$$\|v_h\|_0^2 \leq \frac{(b - a)^2}{4} |v_h|_1^2.$$



Exercise 2.12 Prove Remark 2.4.

Energy estimate

Multiplying both sides of the FD schema $L_h u_i = f_i$ by $u_i \bar{h}_i$ yields

$$-\left((u_i)_{\bar{x}} \right)_{\hat{x}} u_i \bar{h}_i = f_i u_i \bar{h}_i, \quad \forall i = 1, 2, \dots, N.$$

Summing in i gives

$$-\left(((u_h)_{\bar{x}})_{\hat{x}}, u_h \right)_{I_h} = (f_h, u_h)_{I_h}.$$

In virtue of Difference Green Formula and the fact that $u_0 = u_{N+1} = 0$, we have

$$\left((u_h)_{\bar{x}}, (u_h)_{\bar{x}} \right)_{I_h^+} = (f_h, u_h)_{I_h}.$$

Thus

$$\| (u_h)_{\bar{x}} \|_0^2 \leq (f_h, u_h)_{I_h} \leq \| f_h \|_0 \| u_h \|_0 \leq c \| f_h \|_0 \| (u_h)_{\bar{x}} \|_0.$$



That is

$$\|(u_h)_{\bar{x}}\|_0 \leq c \|f_h\|_0.$$

Generalization to

$$-(p(x)u')'(x) + q(x)u(x) = f(x), \quad x \in \Omega,$$

where $p(x) > p_0 > 0, q(x) \geq 0, \forall x \in \Omega$.

Multiplying both sides of the FD schema $L_h u_i = f_i$ by $u_i \bar{h}_i$ yields

$$-(p_{i-1/2}(u_i)_{\bar{x}})_{\hat{x}} u_i \bar{h}_i + q_i u_i u_i \bar{h}_i = f_i u_i \bar{h}_i, \quad \forall i = 1, 2, \dots, N.$$

Summing in i gives

$$-\left((p_h(u_h)_{\bar{x}})_{\hat{x}}, u_h \right)_{I_h} + (q_h u_h, u_h)_{I_h} = (f_h, u_h)_{I_h}.$$

In virtue of Difference Green Formula and the fact that $u_0 = u_{N+1} = 0$, we have

$$(p_h(u_h)_{\bar{x}}, (u_h)_{\bar{x}})_{I_h^+} + (q_h u_h, u_h)_{I_h} = (f_h, u_h)_{I_h}.$$



Thus

$$\|p_0(u_h)_{\bar{x}}\|_0^2 \leq (f_h, u_h)_{I_h} \leq \|f_h\|_0 \|u_h\|_0 \leq c \|f_h\|_0 \|(u_h)_{\bar{x}}\|_0.$$

Error estimate

Exercise 2.13 Derive an estimate for the truncation error of the center schema in case of non-uniform mesh and presence of p .

Error function $e_i = u(x_i) - u_i$ satisfies

$$L_h e_h = R_h, \quad e_0 = e_{N+1} = 0.$$

Therefore

$$\|(e_h)_{\bar{x}}\|_0 \leq c \|R_h\|_0 \leq ch^2.$$



Other methods to establish the stability: In the case of uniform mesh, the scheme can be written under the matrix form

$$A_h u_h = f_h.$$

where A_h is defined in (3). Thus

$$\|u_h\|_2 \leq \|A_h^{-1}\|_2 \|f_h\|_2.$$

A_h is symmetric positive definite, A_h^{-1} is too. Therefore, $\|A_h^{-1}\|_2$ is the maximum eigenvalue of A_h^{-1} , which is the smallest eigenvalue of A_h . It is well-known [Chapter 2, Finite Difference Methods for Ordinary and Partial Differential Equations, LeVeque 1955] that The eigenvalues of A_h are

$$\lambda_p = -4(N+1)^2 \sin^2 \frac{p\pi}{2(N+1)}, \quad p = 1, 2, \dots, N.$$

The eigenvalue corresponding to λ_p is

$$\phi^p = (\phi_1^p, \dots, \phi_N^p)^T, \quad \phi_j^p = \sin \frac{p\pi j}{N+1}, \quad j = 1, \dots, N.$$



The smallest eigenvalue, i.e., λ_1 , behaves (Taylor formula)

$$\lambda_1 = \pi^2 + O(h^2).$$

Consequently,

$$\|u_h\|_2 \leq c\pi^{-2}\|f_h\|_2.$$



Numerical example

Test for an exact solution $u(x) = \sin(x)$:

$$\begin{aligned}-u''(x) &= \sin(x), \quad \forall x \in (0, 2\pi), \\ u(0) &= u(2\pi) = 0.\end{aligned}$$

Numerical solutions for several N .

N	Errors in L^∞ -norm
10	3.191592653460384E-002
20	8.265416966228623E-003
40	2.058706764533680E-003
80	5.142004781495402E-004

Table 2: Errors as a function of N .



Dirac function and Green function

Let $\delta(x)$ be the Dirac function, defined by $\int_{-\infty}^{\infty} \delta(x) v(x) dx = v(0), \forall v \in \mathcal{S}(R)$.

$\delta(x)$ can be viewed as limit $\lim_{\varepsilon \rightarrow 0} \delta_\varepsilon(x)$ with $\delta_\varepsilon(x)$ being $\begin{cases} (\varepsilon+x)/\varepsilon^2 & [-\varepsilon, 0] \\ (\varepsilon-x)/\varepsilon^2 & (0, \varepsilon] \\ 0 & \text{others.} \end{cases}$

$$\cdot \int_{-\infty}^{\infty} \delta(x) dx = \lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{\infty} \delta_\varepsilon(x) dx = 1.$$

Note that $\int_{-\infty}^{\infty} |\delta_\varepsilon(x)|^p dx = \frac{2}{p+1} \varepsilon^{1-p} \rightarrow \infty$ as $\varepsilon \rightarrow 0$, $p > 1$.

• $\delta(x) = H'(x)$ in the distribution sense, where $H(x)$ is the Heaviside function:

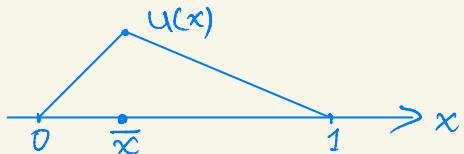
$$H(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 1. \end{cases} \quad H(x) = \lim_{\varepsilon \rightarrow 0} H_\varepsilon(x), \quad H_\varepsilon(x) = \int_{-\infty}^x \delta_\varepsilon(s) ds.$$

Now consider $\begin{cases} -u''(x) = \delta(x-\bar{x}), & x, \bar{x} \in (0, 1), \\ u(0) = u(1) = 0. \end{cases}$

Its solution is called Green function, denoted by $G(x, \bar{x})$.

* How to compute $G(x, \bar{x})$? clearly

$$u(x) = \begin{cases} c \frac{x}{\bar{x}}, & x \in (0, \bar{x}] \\ c \frac{1-x}{1-\bar{x}}, & x \in [\bar{x}, 1). \end{cases}$$



Note that $\int_{\bar{x}-\varepsilon}^{\bar{x}+\varepsilon} u''(x) dx = \int_{\bar{x}-\varepsilon}^{\bar{x}+\varepsilon} \delta(x-\bar{x}) dx = 1$. We have

$$u'(\bar{x}-\varepsilon) - u'(\bar{x}+\varepsilon) = 1.$$

Therefore $\frac{c}{\bar{x}} + \frac{c}{1-\bar{x}} = 1$, $c = \bar{x}(1-\bar{x})$.

Thus give $G(x; \bar{x}) = u(x) = \begin{cases} (1-\bar{x})x, & x \in (0, \bar{x}] \\ \bar{x}(1-x), & x \in [\bar{x}, 1). \end{cases}$

It is readily seen:

1) If $f = \sum_{i=1}^N f_i \delta(x - x_i)$. Then the corresponding solution

$$u(x) = \sum_{i=1}^N f_i G(x; x_i).$$

2) Generally, since $f(x) = \int_0^1 f(\bar{x}) \delta(\bar{x}-x) d\bar{x}$, the solution

$$u(x) = \int_0^1 f(\bar{x}) G(x; \bar{x}) d\bar{x}.$$



Return back to finding $A^{-1} \triangleq (B_1, \dots, B_N)$, s.t.

$AB_j = e_j$, e_j being the j -column of I (identity matrix).

Note that e_j can be viewed as the discrete version of $hS(x-x_j)$, or as a grid function which is 1 in $(x_j - \frac{h}{2}, x_j + \frac{h}{2})$, 0 outside.

So, it is expected that the corresponding solution

$$B_j^{(i)} = hG(x_i; x_j) = \begin{cases} h(1-x_j)x_i & 1 \leq i \leq j \\ h(x_i - x_j)x_j & j \leq i \leq N. \end{cases}$$

(CHECK!)

Return back to the schema $AU = F$, i.e. $U = A^{-1}F$.

We have $\|U\|_\infty \leq \|A^{-1}\|_\infty \|F\|_\infty = \max_{1 \leq i \leq N} \sum_{j=1}^N |B_j^{(i)}| \|F\|_\infty \leq Nh \|F\|_\infty \leq \|F\|_\infty$.
 $(|B_j^{(i)}| \leq h, \forall i, j)$



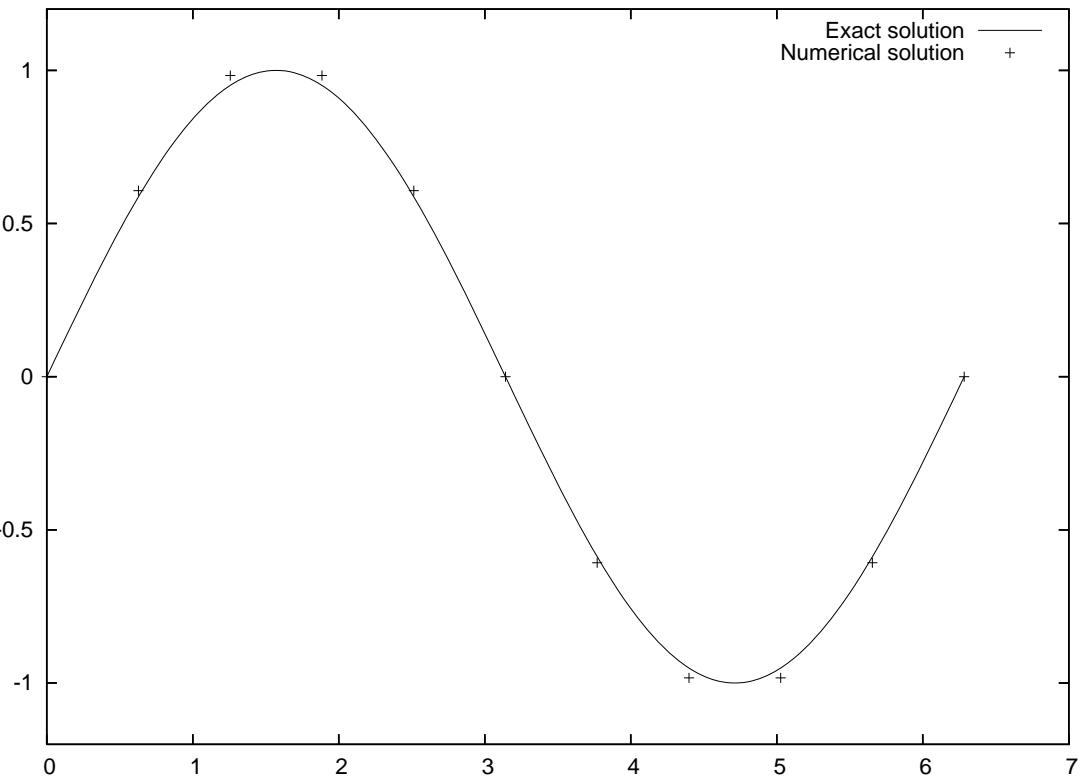


Figure 4: Numerical solution for $N = 10$.

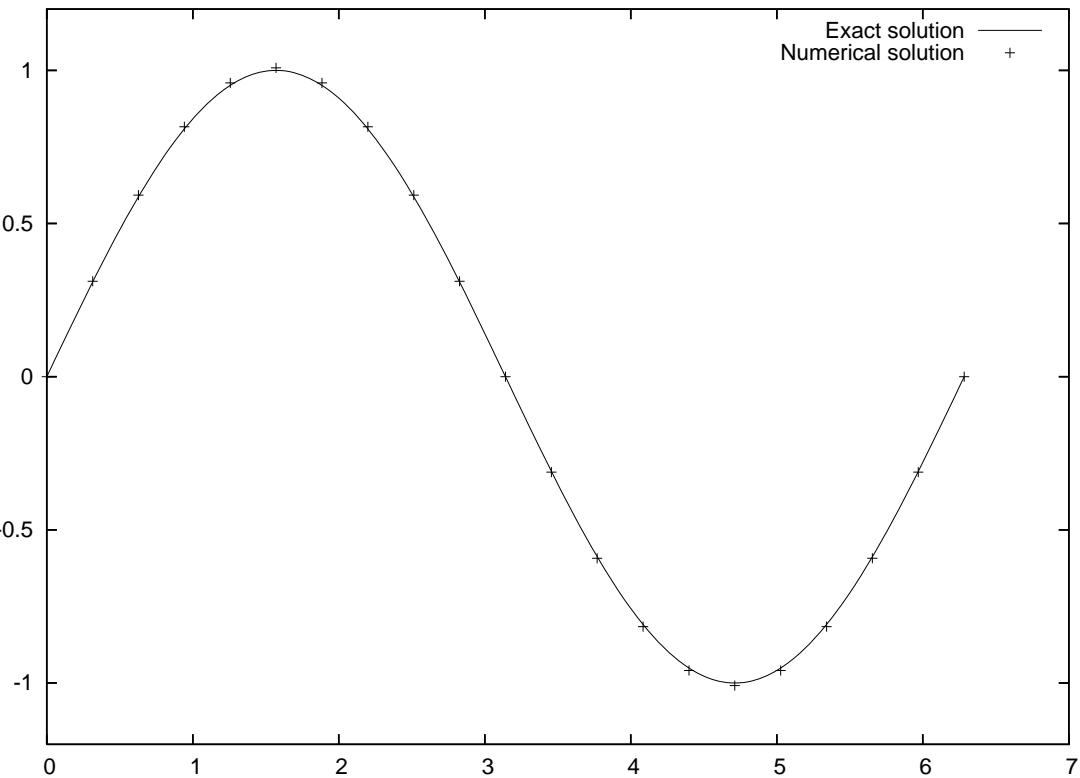


Figure 5: Numerical solution for $N = 20$.

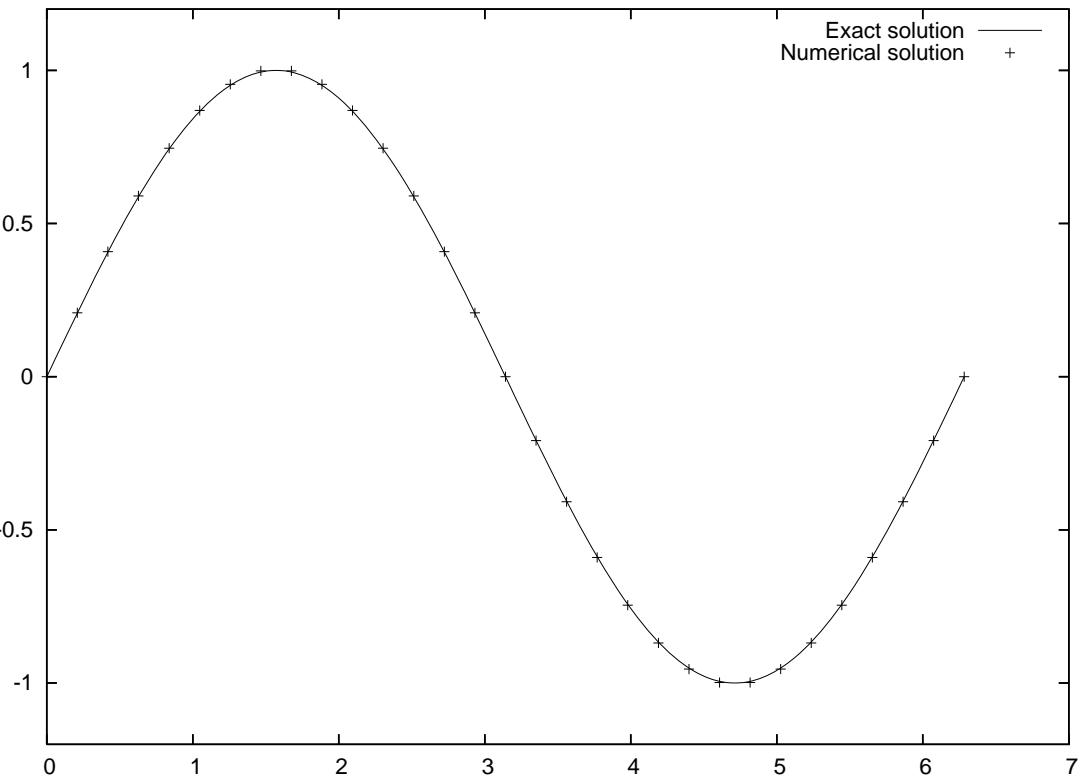


Figure 6: Numerical solution for $N = 30$.

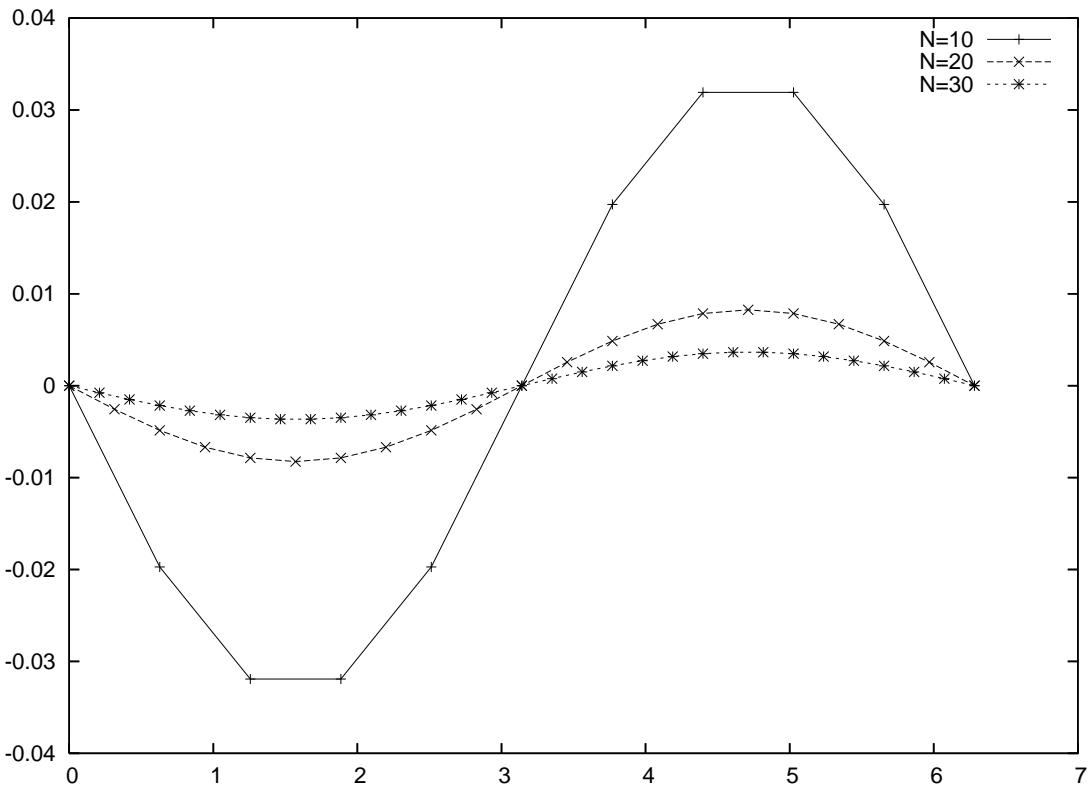


Figure 7: Errors for different $N = 10, 20, 30$.

Elliptic equations with non-homogeneous conditions

Let's consider the following problem

$$\begin{cases} Lu = f, \quad \forall x \in (a, b) \\ u(a) = \alpha, u(b) = \beta. \end{cases}$$

Let $u_h = \bar{u}_h + \tilde{u}_h$, where \bar{u}_h satisfies $\bar{u}_0 = \alpha, \bar{u}_N = \beta$. For example,

$$\bar{u}_i = \frac{\beta - \alpha}{b - a}(x_i - a) + \alpha, \quad \forall i = 0, 1, \dots, N.$$

Then

$$\begin{cases} L_h \tilde{u}_i = f_i - L_h \bar{u}_i, \quad \forall i \in I_h \\ \tilde{u}_0 = \tilde{u}_N = 0. \end{cases}$$

Green formula + Cauchy and Poincaré inequalities \Rightarrow

$$\|(\tilde{u}_h)_{\bar{x}}\|_0 \leq \|f_h\|_0 + c\|\bar{u}_h\|_1.$$



It is readily seen that

$$\|\bar{u}_h\|_1 \leq \sqrt{(b-a) + (b-a)^{-1}}(|\alpha| + |\beta|).$$

As a consequence, we have

$$\|\tilde{u}_h\|_1 \leq c\|(\tilde{u}_h)_{\bar{x}}\|_0 \leq c\|f_h\|_0 + c(|\alpha| + |\beta|).$$

Finally, the triangle inequality gives

$$\|u_h\|_1 \leq c(\|f_h\|_0 + |\alpha| + |\beta|).$$

Exercise 2.14 Consider the elliptic problem

$$\begin{aligned} -u_{xx} + u_x + u &= f, \quad \forall x \in (a, b), \\ u(a) = u(b) &= 0, \end{aligned}$$



and its finite difference schema

$$\begin{aligned} -\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + \frac{u_{i+1} - u_{i-1}}{2h} + u_i &= f_i, \quad \forall i = 1, \dots, N-1, \\ u_0 = u_N &= 0, \end{aligned}$$

in an uniform mesh $\{x_i\}_{i=0}^N, x_i = a + ih, h = (b - a)/N$.

- 1) Derive an estimate for the truncation error;
- 2) Establish an a priori estimate for $\|u_h\|_1$;
- 3) Prove the existence and uniqueness of the solution of the finite difference schema;
- 4) Derive an error estimate for $\|e_h\|_1$, where $e_i = u(x_i) - u_i$.

Exercise 2.15 Consider the elliptic problem

$$\begin{aligned} -u_{xx} &= f, \quad \forall x \in (a, b), \\ u(a) &= 0, \quad u'(b) = \beta, \end{aligned}$$



and its finite difference schema

$$\begin{aligned}-\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} &= f_i, \quad \forall i = 1, \dots, N-1, \\ u_0 &= 0, \\ \frac{u_N - u_{N-1}}{h} &= \beta\end{aligned}$$

in an uniform mesh $\{x_i\}_{i=0}^N, x_i = a + ih, h = (b - a)/N$.

1) Derive an estimate for the truncation errors:

$$R_i^{(1)} = L_h u(x_i) - [Lu](x_i), \quad R^{(2)} = \frac{u_N - u_{N-1}}{h} - u'(b).$$

2) Rewrite the discrete problem under matrix form;

3) Establish an a priori estimate for $\|u_h\|_1$;

4) Derive an error estimate for $\|e_h\|_1$, where $e_i = u(x_i) - u_i$.



Elliptic equations in 2D

- 2D grid

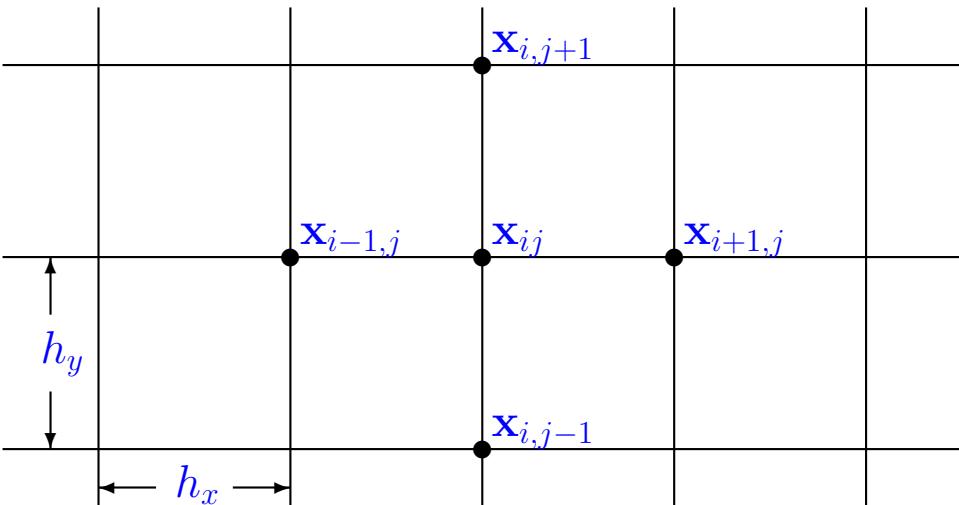


Figure: 2D mesh.

- 2D problem: $\Omega := (a, b)^2$

$$\begin{aligned} Lu(\mathbf{x}) &= f(\mathbf{x}), \quad \forall \mathbf{x} \in \Omega \\ u(\mathbf{x}) &= 0, \quad \forall \mathbf{x} \in \partial\Omega, \end{aligned}$$

where $Lu = -\Delta u = -\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right)$.

- Centered schema: second order accuracy

Let

$$L_h u_{ij} = -\frac{u_{i+1,j} - 2u_{ij} + u_{i-1,j}}{h_x^2} - \frac{u_{i,j+1} - 2u_{ij} + u_{i,j-1}}{h_y^2}.$$

Then the truncation error

$$R_{ij} = L_h u(x_{ij}) - [Lu](x_{ij}) = O(h_x^2 + h_y^2).$$

Error estimate follows the standard procedure (consistency + stability).



Question: how to construct center schema at the grid points close to the boundary in general domains?

- Higher order schemes

$$R_{ij} = L_h u(\mathbf{x}_{ij}) - [Lu](\mathbf{x}_{ij}) = -\frac{1}{12} \left[h_x^2 \frac{\partial^4 u}{\partial x^4} + h_y^2 \frac{\partial^4 u}{\partial y^4} \right] (\mathbf{x}_{ij}) + O(h_x^4 + h_y^4).$$

Furthermore

$$\begin{aligned} h_x^2 \frac{\partial^4 u}{\partial x^4} + h_y^2 \frac{\partial^4 u}{\partial y^4} &= \left(h_x^2 \frac{\partial^2}{\partial x^2} + h_y^2 \frac{\partial^2}{\partial y^2} \right) \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) - (h_x^2 + h_y^2) \frac{\partial^4 u}{\partial x^2 \partial y^2}. \\ &= - \left(h_x^2 \frac{\partial^2}{\partial x^2} + h_y^2 \frac{\partial^2}{\partial y^2} \right) f - (h_x^2 + h_y^2) \frac{\partial^4 u}{\partial x^2 \partial y^2}. \end{aligned}$$

$$\frac{\partial^4 u}{\partial x^2 \partial y^2}(\mathbf{x}_{ij}) = \frac{u_{xx}(x_i, y_{j+1}) - 2u_{xx}(x_i, y_j) + u_{xx}(x_i, y_{j-1})}{h_y^2} + O(h_y^2).$$



Then

$$u_{xx}(x_i, y_{j+1}) = \frac{u(x_{i+1}, y_{j+1}) - 2u(x_i, y_{j+1}) + u(x_{i-1}, y_{j+1})}{h_x^2} \\ + \frac{h_x^2}{12} \frac{\partial^4 u}{\partial x^4}(x_i, y_{j+1}) + O(h_x^4).$$

Similarly for $u_{xx}(x_i, y_j)$ and $u_{xx}(x_i, y_{j-1})$. Finally

$$\frac{\partial^4 u(\mathbf{x}_{ij})}{\partial x^2 \partial y^2} = \frac{1}{h_x^2 h_y^2} [u(x_{i+1}, y_{j+1}) - 2u(x_i, y_{j+1}) + u(x_{i-1}, y_{j+1}) \\ - 2u(x_{i+1}, y_j) + 4u(x_i, y_j) - 2u(x_{i-1}, y_j) \\ + u(x_{i+1}, y_{j-1}) - 2u(x_i, y_{j-1}) + u(x_{i-1}, y_{j-1})] \\ + \frac{h_x^2}{12 h_y^2} \left[\frac{\partial^4 u}{\partial x^4}(x_i, y_{j+1}) - 2 \frac{\partial^4 u}{\partial x^4}(x_i, y_j) + \frac{\partial^4 u}{\partial x^4}(x_i, y_{j-1}) \right] \\ + O(h_x^4/h_y^2) + O(h_y^2).$$



Now we define the modified schema

$$\begin{aligned}\bar{L}_h u_{ij} &= \bar{f}(\mathbf{x}_{ij}), \quad \forall(i, j) \text{ s.t. } \mathbf{x}_{ij} \in \Omega, \\ u_{ij} &= 0, \quad \forall(i, j) \text{ s.t. } \mathbf{x}_{ij} \in \partial\Omega,\end{aligned}$$

where the finite difference operator \bar{L}_h is defined by

$$\begin{aligned}\bar{L}_h u_{ij} = L_h u_{ij} - \frac{h_x^2 + h_y^2}{12h_x^2 h_y^2} [&u_{i+1,j+1} - 2u_{i,j+1} + u_{i-1,j+1} \\ &- 2u_{i+1,j} + 4u_{i,j} - 2u_{i-1,j} \\ &+ u_{i+1,j-1} - 2u_{i,j-1} + u_{i-1,j-1}],\end{aligned}$$

$\bar{f}(\mathbf{x}_{ij})$ is given by

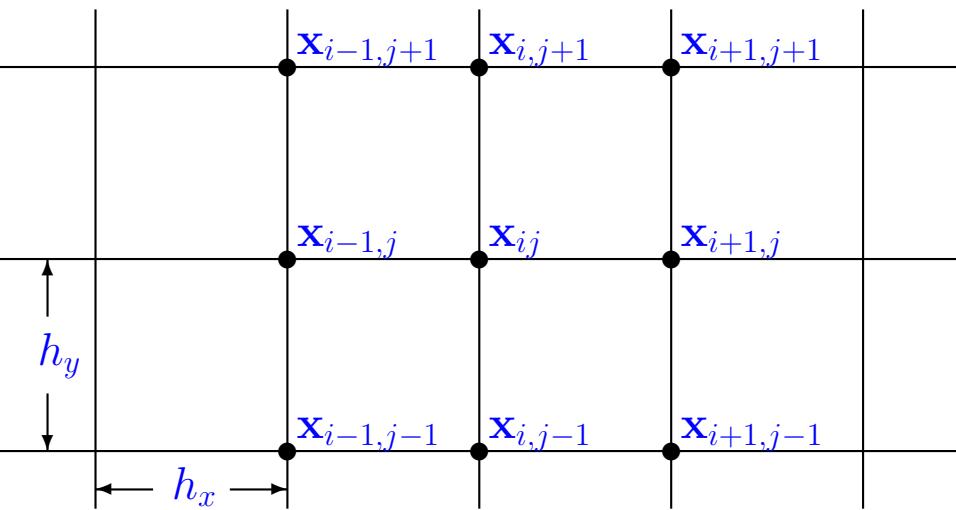
$$\bar{f}(\mathbf{x}_{ij}) = f(\mathbf{x}_{ij}) + \frac{1}{12} \left(h_x^2 \frac{\partial^2 f}{\partial x^2} + h_y^2 \frac{\partial^2 f}{\partial y^2} \right)(\mathbf{x}_{ij}).$$

Then the truncation error has the following estimate:

$$\bar{R}_{ij} := \bar{L}_h(\mathbf{x}_{ij}) - \bar{f}(\mathbf{x}_{ij}) = O(h_x^4 + h_y^4), \quad \text{if } h_x = O(h_y).$$



⇒ 9-point schema!



Exercise 3.1 Derive an estimate for the truncation error of the 9-point schema.



Parabolic equations

Consider the heat conduction problem

$$\begin{aligned} u_t - u_{xx} &= f, \quad \forall t \in (0, T], \forall x \in (a, b), \\ u(x, 0) &= u_0(x), \quad \forall x \in (a, b), \\ u(a, t) = u(b, t) &= 0, \quad \forall t \in (0, T]. \end{aligned}$$

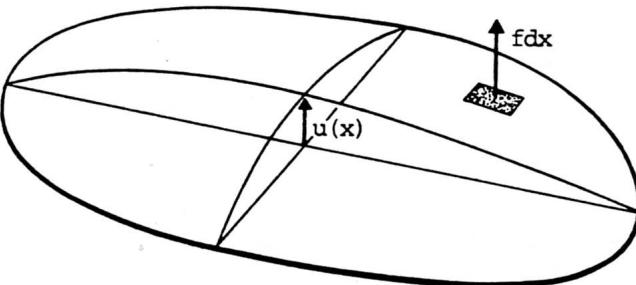


Figure 8: Displacement of an elastic membrane.



1. Forward Euler/centered schema

$$\frac{u_i^{n+1} - u_i^n}{k} - \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{h^2} = f_i^{n+1}, \quad n = 0, 1, \dots, M; i = 1, \dots, N-1,$$

where u_i^n is an approximation of $u(x_i, t^n)$.

- 1+1-dimensional grid

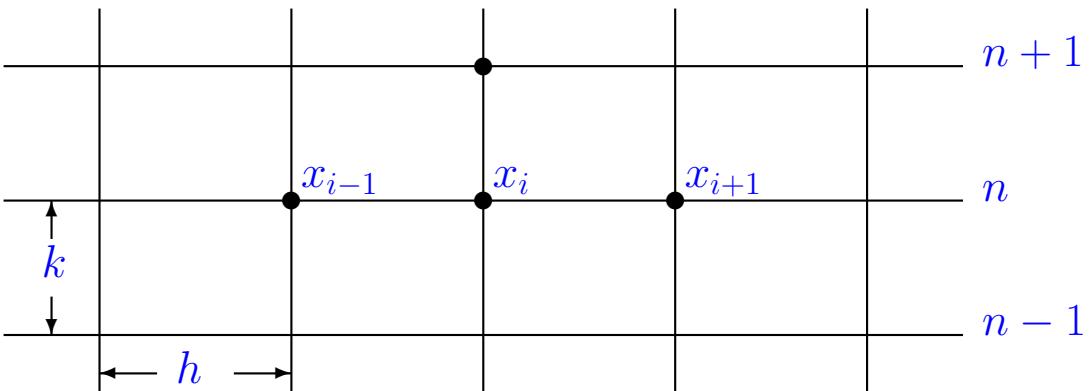


Figure: 1+1D mesh.

- Truncation error

Let

$$Lu = u_t - u_{xx}, \quad L_h u_i^n = \frac{u_i^{n+1} - u_i^n}{k} - \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{h^2}.$$

Then

$$R_i^n := L_h u(x_i, t^n) - [Lu](x_i, t^n) = \frac{1}{2} k u_{tt} - \frac{1}{12} h^2 u_{xxxx} + \dots = O(k + h^2). \quad (6)$$

Exercise 3.2 Prove (6).

- Stability ($f = 0$)

$$\begin{aligned} u_i^{n+1} &= u_i^n + \tau(u_{i+1}^n - 2u_i^n + u_{i-1}^n) \\ &= (1 - 2\tau)u_i^n + \tau u_{i+1}^n + \tau u_{i-1}^n. \end{aligned}$$



where $\tau = \frac{k}{h^2}$.

Suppose $\tau \leq 1/2$ and denote $U^n = \max_i |u_i^n|$, then

$$U^{n+1} \leq (1 - 2\tau)U^n + \tau U^n + \tau U^n = U^n.$$

Thus

$$U^n \leq U^0, \quad \forall n = 1, 2, \dots.$$

Stability condition

- For $\tau > 1/2$ the numerical solution is not bounded!
- A strict condition for mesh refinement:

$$\tau = k/h^2 \leq 1/2$$

i.e. $k \leq h^2/2$.

- Error estimate



Let pointwise error $e_i^n = u(x_i, t^n) - u_i^n$. Then

$$e_i^{n+1} = (1 - 2\tau)e_i^n + \tau e_{i+1}^n + \tau e_{i-1}^n + kR_i^n.$$

Suppose $\tau \leq 1/2$ and denote $E^n = \max_i |e_i^n|$, then

$$E^{n+1} \leq E^n + k \max_i |R_i^n|.$$

If $E^0 = 0$, we have

$$E^n \leq nk \max_i |R_i^n| \leq TO(k + h^2), \quad \forall n = 1, 2, \dots.$$

- Goal: improved accuracy

⇒ smaller h

⇒ much smaller time step size k

⇒ more unknowns, more work

⇒ larger rounding errors



\Rightarrow impaired accuracy

- Numerical experiments

Consider the heat conduction problem

$$\begin{aligned} u_t - u_{xx} &= 0, \quad \forall t \in (0, T], \forall x \in (0, 5), \\ u(x, 0) &= \sin(x), \quad \forall x \in (0, 5), \\ u(0, t) &= 0, u(5, t) = \sin(5), \quad \forall t \in (0, T]. \end{aligned}$$

Resolution: $N = 100, h = 5/100.$

1. Stable calculation: $k = \frac{1}{2}h^2 = 0.00125$ such that $\tau = \frac{k}{h^2} = \frac{1}{2}.$





2. Unstable calculation: $k = 0.001275$ such that $\tau = \frac{k}{h^2} > \frac{1}{2}$.



2. Backward Euler/centered schema

$$\frac{u_i^{n+1} - u_i^n}{k} - \frac{u_{i+1}^{n+1} - 2u_i^{n+1} + u_{i-1}^{n+1}}{h^2} = 0, \quad n = 0, 1, \dots, M; i = 1, \dots, N-1.$$

\Rightarrow

$$u_i^{n+1} - \tau(u_{i+1}^{n+1} - 2u_i^{n+1} + u_{i-1}^{n+1}) = u_i^n.$$

\Rightarrow

$$A\mathbf{u}^{n+1} = \mathbf{u}^n,$$

where A is a tridiagonal matrix.

Explicit and Implicit Schemes



★ Explicit

- Difference schema allows the solution of one unknown value u_i at a time.
- Little work/unknown, but may be unstable.

★ Implicit

- Several unknown values must be solved simultaneously.
 - More work/unknown, but more robust than explicit schemes (less severe step size limitations)
 - smaller amount of total work
- ★ Combination of both methods in one problem possible.



Exercise 3.3 Consider the transport-diffusion problem

$$\begin{aligned} u_t - u_{xx} + vu_x &= 0, \quad \forall x \in (a, b), t \in (0, T) \\ u(a, t) = u(b, t) &= 0, \quad t \in (0, T) \\ u(x, 0) &= u_0(x), \quad \forall x \in (a, b) \end{aligned}$$

where v is a constant. Derive an estimate for the truncation error of the following schema, and prove that

$$\|u_h^n\|_0 \leq \|u_h^0\|_0, \quad \forall n = 0, 1, \dots,$$

- If $v \geq 0$,

$$\begin{aligned} \frac{u_i^{n+1} - u_i^n}{k} - \frac{u_{i+1}^{n+1} - 2u_i^{n+1} + u_{i-1}^{n+1}}{h^2} + v \frac{u_i^{n+1} - u_{i-1}^{n+1}}{h} &= 0, \quad \forall i = 1, \dots, N-1, \\ u_0^{n+1} &= u_N^{n+1} = 0, \\ u^0 &= u_0, \end{aligned}$$



- If $v \leq 0$,

$$\begin{aligned} \frac{u_i^{n+1} - u_i^n}{k} - \frac{u_{i+1}^{n+1} - 2u_i^{n+1} + u_{i-1}^{n+1}}{h^2} + v \frac{u_{i+1}^{n+1} - u_i^{n+1}}{h} &= 0, \quad \forall i = 1, \dots, N-1, \\ u_0^{n+1} = u_N^{n+1} &= 0, \\ u^0 &= u_0, \end{aligned}$$

in an uniform mesh $\{x_i\}_{i=0}^N, x_i = a + ih, h = (b-a)/N, \{t^n\}_{n=0}^M, t^n = nk, k = T/M$.

